Defense Technical Information Center
Compilation Part Notice

ADP012691

TITLE: Consistency-Based Fault Isolation for Uncertain Systems with Applications to Quantitative Dynamic Models

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: Thirteenth International Workshop on Principles of Diagnosis [DX-2002]

To order the complete compilation report, use: ADA405380

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP012686 thru ADP012711

Consistency-Based Fault Isolation for Uncertain Systems with Applications to Quantitative Dynamic Models

Colin N. Jones [1] and Gregory W. Bond [2] and Peter D. Lawrence[3]

Abstract. This paper presents the Probabilistic General Diagnostic Engine (PGDE), a novel method of offline consistency-based fault isolation. Many existing proposals require qualitative logic models for consistency-based diagnosis due to their ability to speed the search for conflict sets through the use of an ATMS. However, for many applications, quantitative dynamic models are preferred or already available. The key strength of the PGDE is that it allows the use of any modelling language for which an appropriate calculation engine can be written. It also offers graceful degradation in the presence of uncertainty, commonly caused by noise or modelling errors. Finally, given perfect knowledge, it can be shown that the PGDE computes the same result as existing consistency-based diagnosis methods. To demonstrate the performance of the algorithm, we have used a quantitative dynamic model of the fluid power circuit of a single-degree of freedom hydraulic test bench and developed an appropriate calculation engine for computing consistency between measured values and predicted results. Various failures were generated on the physical test bench and the PGDE isolated the faults with approximately 85% accuracy.

1 INTRODUCTION

Consistency-based diagnosis has at its heart the search for a subset of the full model such that predictions made using the subset are consistent with sensor measurements. This search space is exponential in the number of model components and so a great deal of attention has been given to developing efficient algorithms. Much progress has been made by utilizing the properties of propositional logic and qualitative models ([10, 8, 1] to name a few) but the problems associated with more complex dynamic systems have still to be solved in general. The Probabilistic General Diagnostic Engine (PGDE) addresses some of these issues in a general framework that applies to any model for which an appropriate "consistency measure" can be formulated.

There are many devices for which quantitative dynamic models either already exist or whose behavior can best be described by a set of differential equations. The cost of developing qualitative models exclusively for the purpose of diagnosis is prohibitive, thus making the adaptation of qualitative methods to quantitative dynamic models an important topic. Models of this type present two new challenges to the diagnostician: First, quantitative dynamic models require comparison of sets of signals to determine consistency. Due to noise and modelling errors, it can be difficult to represent the results of these comparisons by the discrete values typically used in qualitative methods. Second, the nature of dynamic systems is that they often have states which are not directly measurable. When the model is simulated using only the equations from a few components, it is often the case that many of the states will become unknown. If no conflict is observed, we reason that a possible diagnosis has been identified, however, it is impossible to know if there would have been a conflict if these states had been known. As a result, the underconstrained nature of dynamic systems reduces the resolution of fault isolation procedures and this must be taken into account in any diagnostic method dealing with these models.

The PGDE algorithm attempts to deal with these difficulties by maintaining a belief distribution for each possible diagnosis. Since these distributions are not limited to discrete-valued consistency measures, the PGDE is able to more accurately interpret intermediate non-boolean consistency assessments. They are also updated throughout the duration of the diagnostic procedure, and conclusions about the consistency of sets of components with observations are not drawn until sufficient information has been processed. In Section 2, the proposed algorithm is laid out in a step-by-step fashion, including consideration of its computational complexity in Section 2.5. Next, Section 3 presents a non-trivial example hydraulic circuit and summarizes some diagnostic results obtained by the PGDE. Finally, the paper closes with a discussion of conclusions and future directions of research in Section 4.

2 PGDE ALGORITHM

The model used in a consistency-based algorithm is a set of constraints on the signals passing through the system. A failure can be declared when these signals are inconsistent with the constraints. The goal of the algorithm is then to locate a subset of these constraints, which when removed from the model, restore consistency between the predicted and observed behavior. This process can proceed in an iterative manner, selecting a set of constraints to remove and simulating the system until a feasible set is found.

We begin by defining the system as in [7]:

Definition 1 *A system is a triple (SD,COMPS,OBS) where:*

1. *the components (COMPS) are a finite set of constants*
2. *the system description (SD) is a set of constraints*
3. *the observations (OBS) are measurements of the physical device*

There is no requirement that there be a one-to-one mapping from components to constraints and so a partition $\{SD_c\}_{c \in COMPS}$ is defined covering SD such that $\bigcup_{c \in COMPS} SD_c = SD$ and $SD_{c_i} \bigcap SD_{c_j} = \emptyset \ \forall c_i \neq c_j$. The set of all possible failures is

---

given by the power set of $COMPS$ and for each element $\Delta \subseteq P(COMPS)$, define $SD_\Delta = \bigcup_{c \in \Delta} SD_c$. This allows the definition of components which contain large numbers of constraints or complex behaviors as well as hierarchies of components. The cardinality of a set of constraints $X \subseteq SD$ is written as $|X|$; it is a system-dependent real number, representing the notion of how "large" the set X is when compared to SD.

Reiter's original work [7] relies on a 'theorem prover', $\mathsf{TP}(SD, \mathcal{D}(\Delta, COMPS \backslash \Delta), OBS)$, which returns true if the partial model containing only the constraints in the complement of SD_Δ, $(SD_\Delta)^c$, is consistent with the observations OBS and false otherwise; consistency implying that the components Δ are a possible diagnosis. Here the theorem prover is redefined to return a continuous measure of how consistent the constraints $(SD_\Delta)^c$ are with the observations OBS. It is possible that the system defined by $(SD_\Delta)^c$ with OBS as inputs may be underconstrained. Thus, for some of the constraints in $(SD_\Delta)^c$, it is impossible to verify if they have, or have not, been violated. If this system is consistent then it is not valid to say that Δ is a diagnosis as the faults might have been in the constraints that could not be tested. This situation is very common in dynamic systems with state as they are inherently underconstrained [4]. To deal with this, the constraints which were used during the simulation of $(SD_\Delta)^c$ are returned by $\mathsf{TP}(\cdot)$ as defined below.

Definition 2 *Let $\Delta \in P(COMPS)$. Define the function* $\mathsf{TP}(\cdot, \cdot)$: $SD \times OBS \rightarrow \mathbb{R} \times SD$ *as:*

$$(\mu_\Delta, A_\Delta) = \mathsf{TP}((SD_\Delta)^c, OBS)$$

Where:

- $\mu_\Delta \in [0,1]$, 1 *implies constraints $(SD_\Delta)^c$ are consistent with the observations OBS, and 0 implies inconsistency*

- $A_\Delta \subseteq (SD_\Delta)^c$ *are the constraints which $\mathsf{TP}(\cdot)$ had sufficient information to apply during the calculation of μ_Δ*

Two belief distributions over the states $\{true, false, unknown\}$ are maintained for each element $\Delta \in P(COMPS)$. These are represented by the probability mass functions $B_{\mathsf{D},\Delta}(x)$ and $B_{\mathsf{IC},\Delta}(x)$ with domains $\{true, false, unknown\}$. $B_{\mathsf{D},\Delta}(true)$ is the belief that the evidence, provided by calls to $\mathsf{TP}(\cdot)$, shows that Δ is a diagnosis. $B_{\mathsf{D},\Delta}(false)$ is the belief that the evidence does not show that Δ is a diagnosis. It does not mean that the evidence *does* show that Δ is *not* a diagnosis as consistency can only incriminate components, it cannot exonerate them [7]. Finally, $B_{\mathsf{D},\Delta}(unknown)$ is the probability that it is unknown what the evidence shows, or that there is no evidence. If $\mu_\Delta = 0$ then at least one component of Δ^c must be faulty and we call Δ^c a conflict set [7] and Δ an inverse conflict. $B_{\mathsf{IC},\Delta}(true)$ is the belief that the evidence shows that Δ is an inverse conflict, $B_{\mathsf{IC},\Delta}(false)$ that it doesn't and $B_{\mathsf{IC},\Delta}(unknown)$ that the evidence is unclear.

Initially, all the beliefs are 100% *unknown* ($B_{\mathsf{D},\Delta}(x) = B_{\mathsf{IC},\Delta}(x) = \{0.0, 0.0, 1.0\}$). In each iteration, a call is made to $\mathsf{TP}(\cdot)$ to check if a new set of constraints $(SD_\Delta)^c$, is consistent with the observations, OBS. The distributions are then updated to reflect the simulator's certainty in the consistency of each set of components, again with the observations. In this way, the diagnostic engine determines the components that are most likely to be faulty, as well as a measure of its confidence in these decisions.

A block diagram of the PGDE is shown in Figure 1. The following sections deal with each stage of the algorithm in detail in the order:

updating the beliefs (steps 3 and 4), choosing a new set to test for consistency via $\mathsf{TP}(\cdot)$ (step 1), deciding when to stop and interpreting the final belief distributions (steps 5 and 6).

2.1 Belief update

Once a possible diagnosis, Δ, has been selected, $\mathsf{TP}(\cdot)$ is used to find the consistency measure, μ_Δ, and the constraints which were used to compute it, A_Δ. The goal is to determine what the consistency measure has shown about each of the subsets of $COMPS$, using A_Δ as a guide. Assuming no fault models, two properties of constraint systems allow the consistency measure of the set Δ to affect the beliefs of other sets: supersets of diagnoses are diagnoses (removing more constraints will not make the system inconsistent) and subsets of inverse conflicts are inverse conflicts (adding constraints will not make the system consistent). Using these facts, the supersets of Δ are first considered and the information derived from μ_Δ and A_Δ is used to update the beliefs that they are diagnoses ($B_{\mathsf{D},\Delta_P}(x) \; \forall \Delta_P \supseteq \Delta$). Similarly, the beliefs that the subsets are inverse conflicts are also updated ($B_{\mathsf{IC},\Delta_C}(x) \; \forall \Delta_C \subseteq \Delta$).

2.1.1 Update belief in diagnosis

We begin by assuming that $\mu_\Delta = 1$, indicating that the observations are consistent with the constraints $(SD_\Delta)^c$. The goal is to determine to what degree this evidence shows that each set is a diagnosis. The first step is to locate the *base set*, Δ_B, for the set $(SD_\Delta)^c$ as defined below in Definition 3. This is the set with the most components of which none have had any of their constraints used during the calculation of μ_Δ. Referring to Figure 2, in which $\mathsf{TP}((SD_{\{1,2,3\}})^c, OBS)$ was called, the base node is $\Delta_B = \{1, 2, 3, 4\}$. If $\Delta \neq \Delta_B$, then the constraints of at least one component have not been considered due to the assumption that the components in Δ were faulty (in Figure 2 this would be component 3). In essence, $\mathsf{TP}(\cdot)$ cannot distinguish between any set Δ' such that $\Delta \subseteq \Delta' \subseteq \Delta_B$, since whenever the constraints associated with the components in Δ are not considered, neither are those of Δ_B, which implies that $\mu_\Delta = \mu_{\Delta'} = \mu_{\Delta_B}$. This is a limitation of the model and the placement of the sensors; as a result the best the algorithm can do is incriminate Δ_B and inform the user of this sensor deficiency. Because the consistency measure would be the same for all of the sets Δ', such that $\Delta \subseteq \Delta' \subseteq \Delta_B$, the sets are marked and ignored in subsequent calls to $\mathsf{TP}(\cdot)$. For certain model types these families of sets can be identified *a priori* and grouped into single components to speed the algorithm [1, 2].

Definition 3 *Let $\Delta \subseteq \Delta_B \subseteq COMPS$. Then Δ_B is the base set for Δ iff*

$$SD_{\Delta_B} \bigcap A_\Delta = \emptyset$$

$$\forall \Delta' \supset \Delta_B, \; SD_{\Delta'} \bigcap A_\Delta \neq \emptyset$$

If the constraints associated with Δ_B are not considered during the call to $\mathsf{TP}(\cdot)$, those in $(A_\Delta)^c \backslash SD_{\Delta_B}$ are not either (in Figure 2 this would be the unshaded sections of components 5 and 6). These are the constraints which were *not* considered that do *not* make up a full component. The question is: Is the lack of conflict during the computation of μ_Δ due to the constraints in SD_{Δ_B}, those in $(A_\Delta)^c \backslash SD_{\Delta_B}$, or some combination of the two? The safest approach would be to say that this evidence can only increase the belief that some set $\Delta' \supseteq \Delta_B$ which covers all of $(A_\Delta)^c$ is a diagnosis ($\Delta' = \{1, 2, 3, 4, 5, 6\}$ in the example). However, if
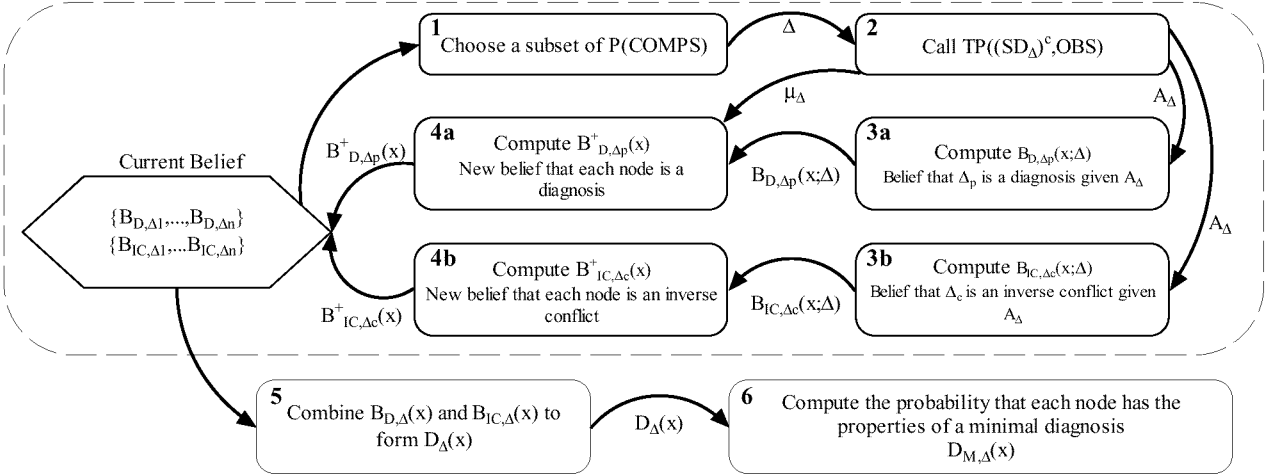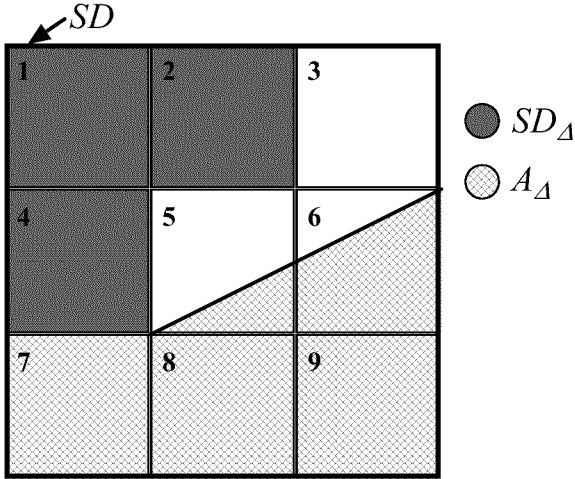
Figure 1. The PGDE Algorithm

Figure 2. Example nine component system

$|(A_\Delta)^c \setminus SD_{\Delta_B}| \ll |SD_{\Delta_B}|$, this would be a very conservative approach, in the sense that a set will never be called a diagnosis if it cannot completely explain the observed behavior, and multiple component failures would be returned more often than they should. In most cases, designing models which reduce the size of $(A_\Delta)^c \setminus SD_{\Delta_B}$ will increase the precision of the diagnosis and so we make the assumption that most modelers will aim for this characteristic and as a result assume that $|(A_\Delta)^c \setminus SD_{\Delta_B}|$ is small compared to $|SD_{\Delta_B}|$.

Under the assumption that the majority of the constraints which were not considered during the computation of μ_Δ belong to Δ_B, this evidence increases the belief that Δ_B is a diagnosis. However, because every superset of a diagnosis is a diagnosis, this evidence also increases the belief that all of the supersets of Δ_B are diagnoses. Therefore for each set $\Delta_P \supseteq \Delta_B$ the probability that the constraints in SD_{Δ_P} can account for the lack of conflict during the computation

of μ_Δ is:

$$P(\Delta_P \text{ is a diagnosis} \mid A_\Delta \wedge \mu_\Delta = 1) \qquad (4)$$
$$= \frac{|(A_\Delta)^c \bigcap SD_{\Delta_P}|}{|(A_\Delta)^c|}$$

Assuming that faults are equally likely to be anywhere in $(A_\Delta)^c$, the probability that they are in SD_{Δ_P} is given by Equation 4, as the proportion of $(A_\Delta)^c$ that is covered by SD_{Δ_P}. If all of, or more than, $(A_\Delta)^c$ is covered, then the probability that the system will be consistent is 100%, by the assumption that $\mu_\Delta = 1.0$.

This probability is computed assuming $\mu_\Delta = 1$, when in fact it may well be less than one. The consistency measure describes our ability to measure how consistent the observations are with the constraints A_Δ. The real components Δ are either consistent or inconsistent with observations and it is only the inability of the model and sensors to perfectly determine which one is true that causes $\mu_\Delta < 1$. Therefore the consistency measure can be interpreted as a probability that the real artifact is consistent or inconsistent and we assume a mapping $\mathcal{PC}(\mu_\Delta)$ to $[0, 1]$ defined by the modeler which represents how probable it is that the real artifact is consistent given μ_Δ.

For each $\Delta_P \supseteq \Delta_B$ we define a belief distribution $B_{D,\Delta_P}(x; \Delta)$ over the states $\{true, false, unknown\}$ which represents the belief that Δ_P is a diagnosis given only the information from calling TP$(\cdot)$ on Δ. The distribution is defined as follows:

$$B_{D,\Delta_P}(true; \Delta)$$
$$= P(\Delta_P \text{ is a diagnosis} \mid A_\Delta \wedge \mu_\Delta = 1) \cdot \mathcal{PC}(\mu_\Delta)$$
$$B_{D,\Delta_P}(false; \Delta)$$
$$= (1 - P(\Delta_P \text{ is a diagnosis} \mid A_\Delta \wedge \mu_\Delta = 1)) \cdot \mathcal{PC}(\mu_\Delta)$$
$$B_{D,\Delta_P}(unknown; \Delta)$$
$$= 1 - \mathcal{PC}(\mu_\Delta) \qquad (5)$$

Equation 5 takes the probability that a set is a diagnosis given A_Δ and that the measure is consistent, and then scales this probability by the certainty that the call to TP$(\cdot)$ returned consistent. This distribution is now combined with the current beliefs using Bayes' Theorem and the Total Probability Theorem.

Let F be the set $\{true, false, unknown\}$. Then the current belief distribution, $B_{\text{D},\Delta_P}(x)$, is updated by the evidence $B_{\text{D},\Delta_P}(x;\Delta)$ to the new belief distribution $B_{\text{D},\Delta_P}^+(x)$:

$$B_{\text{D},\Delta_P}^+(x) = \tag{6}$$
$$\sum_{f_1,f_2 \in F} P(B_{\text{D},\Delta_P}^+(x) \mid B_{\text{D},\Delta_P}(f_1) = 1 \wedge$$
$$B_{\text{D},\Delta_P}(f_2;\Delta) = 1) \cdot B_{\text{D},\Delta_P}(f_1) \cdot B_{\text{D},\Delta_P}(f_2;\Delta)$$

The probabilities $P(B_{\text{D},\Delta_P}^+(x) \mid B_{\text{D},\Delta_P}(f_1) = 1 \wedge B_{\text{D},\Delta_P}(f_2;\Delta) = 1)$ in Equation 6 can be represented by a conditional probability table as shown in Table 1. The first two columns represent f_1 and f_2 respectively and the last three represent x. The values in Table 1 are chosen such that if the current belief is very certain, as defined by the weight of the *unknown* state, then a new distribution which is very uncertain, will not strongly influence the belief, and vice versa. If the new evidence agrees with our current belief, then this belief is strengthened, and if it does not then it is weakened.

Table 1. Conditional Probability Table used to update $B_{\text{D},\Delta_P}(x)$ given $B_{\text{D},\Delta_P}(x;\Delta)$

$$P(B_{\text{D},\Delta_P}^+(x) \mid B_{\text{D},\Delta_P}(f_1) = 1 \wedge B_{\text{D},\Delta_P}(f_2;\Delta) = 1)$$

f_1	f_2	x		
		$True$	$False$	$Unknown$
True	True	1.0	0.0	0.0
True	False	0.5	0.5	0.0
True	Unknown	1.0	0.0	0.0
False	True	0.5	0.5	0.0
False	False	0.0	1.0	0.0
False	Unknown	0.0	1.0	0.0
Unknown	True	1.0	0.0	0.0
Unknown	False	0.0	1.0	0.0
Unknown	Unknown	0.0	0.0	1.0

2.1.2 *Update belief in inverse conflict*

To update the beliefs $B_{\text{IC},\Delta}(x)$, much the same procedure is followed as in the case where the system is consistent, only now the evidence suggests that the considered sets are inverse conflicts rather than diagnoses. As before, the first step is to locate the set Δ_B, but now it is the base set of $(A_\Delta)^c$ ($\Delta_B = \{7, 8, 9\}$ in Figure 2). $(\Delta_B)^c$ is the largest set of components such that all of $(SD_{\Delta_B})^c$ was used to compute μ_Δ and we again assume that $|(SD_{\Delta_B})^c| \gg |A_\Delta \setminus (SD_{\Delta_B})^c|$. The evidence provided by μ_Δ suggests that some of the constraints in $(SD_{\Delta_B})^c$ have been violated. Since adding constraints will not take away the fact that some of these have not been met, every superset of $(SD_{\Delta_B})^c$ also contains broken constraints indicating that every subset, Δ_C, of Δ_B is an inverse conflict. As before, the probability that the set Δ_C is an inverse conflict is:

$$P(\Delta_C \text{ is an inverse conflict} \mid A_\Delta \wedge \mu_\Delta = 0)$$
$$= \frac{|A_\Delta \bigcap (SD_{\Delta_C})^c|}{|A_\Delta|}$$

We assume a mapping $\mathcal{PIC}(\mu_\Delta) \in [0,1]$, defined by the modeler, which represents the probability that the real artifact is inconsistent given μ_Δ. This mapping is then used to compute a distribution, $B_{\text{IC},\Delta_C}(x;\Delta)$, over the states $\{true, false, unknown\}$ which represents the belief that the set Δ_C is an inverse conflict given only

the information from calling $\text{TP}(\cdot)$ on Δ.

$$B_{\text{IC},\Delta_C}(true;\Delta)$$
$$= P(\Delta_C \text{ is an inverse conflict} \mid A_\Delta \wedge \mu_\Delta = 0) \cdot \mathcal{PIC}(\mu_\Delta)$$
$$B_{\text{IC},\Delta_C}(false;\Delta)$$
$$= (1 - P(\Delta_C \text{ is an inverse conflict} \mid A_\Delta \wedge \mu_\Delta = 0))$$
$$\cdot \mathcal{PIC}(\mu_\Delta)$$
$$B_{\text{IC},\Delta_C}(unknown;\Delta)$$
$$= 1 - \mathcal{PIC}(\mu_\Delta)$$

This belief distribution is incorporated into our current belief $B_{\text{IC},\Delta_C}(x)$ in the same manner as discussed in the previous section. The total probability theorem is again used as in Equation 6 to compute the new belief distribution $B_{\text{IC},\Delta_C}^+(x)$ from the old one $B_{\text{IC},\Delta_C}(x)$ and the new evidence $B_{\text{IC},\Delta_C}(x;\Delta)$ using the conditional probabilities in Table 1.

The new evidence provided by the call to $\text{TP}((SD_\Delta)^c, OBS)$ has now been incorporated into the belief distributions $B_{\text{IC},\Delta}(x)$ and $B_{\text{D},\Delta}(x)$ for all subsets Δ of $COMPS$. The next section looks at how to use these belief distributions to choose the next component to pass to $\text{TP}(\cdot)$.

2.2 Next best set

The order in which the subsets of $COMPS$ are tested is crucial to the speed at which the algorithm will find the diagnoses. There are, however, several choices which will produce varying results and so the choice depends largely on knowledge of the system. The following properties can be taken into account when developing a heuristic search strategy:

- Failure rates: choose sets of components with a history of failure
- Expected knowledge gain: choose sets of components which are expected to reduce the unknown portions of the belief distributions the most. (i.e. $B_{\text{D},\Delta}(unknown)$ and $B_{\text{IC},\Delta}(unknown)$). See [5] for a derivation.
- Current belief: choose the supersets and subsets of the set currently most likely to be a minimal diagnosis to isolate a single diagnosis as quickly as possible.
- Principle of Parsimony: choose the sets with the fewest components as they are more likely to be diagnoses.
- Execution time: choose the sets with the most components, as $\text{TP}(\cdot)$ will likely take less time to evaluate systems with fewer constraints.

2.3 Stop conditions

The certainties in the potential diagnoses returned by the PGDE increase monotonically with each iteration [5]. Thus, the maximum certainties are achieved when all subsets of $P(COMPS)$ have been passed to $\text{TP}(\cdot)$ for testing. Since this is likely to take too long, a decision needs to be made about when to stop. As it is when choosing a search algorithm, this decision is mostly heuristic and entirely up to the modeler. Some examples of criteria are listed here:

- A time limit has been reached
- The sum of all of the subsets of $P(COMPS)$'s knowledge has risen above some limit
- The knowledge gained per call to $\text{TP}(\cdot)$ has fallen below some level

- A percentage of the subsets of $COMPS$ have been tested
- At least one minimal diagnosis has been found with some minimum certainty

2.4 Most likely minimal diagnoses

A minimal diagnosis is a diagnosis such that no proper subset of it is also a diagnosis. They are of interest as the Principle of Parsimony [7] states that the diagnoses with the fewest components are the most likely. The minimal diagnoses will have the properties that all of their supersets will be diagnoses and all of their proper subsets will be inverse conflicts. The goal is to determine which sets are most likely to have these properties given the belief distributions $B_{IC,\Delta}(x)$ and $B_{D,\Delta}(x)$.

2.4.1 Combining $B_D(x)$ and $B_{IC}(x)$

The two belief distributions $B_D(x)$ and $B_{IC}(x)$ have been kept separate, as they represent different types of information. In order to compute the most likely minimal diagnoses, all of the information needs to be taken into account and as a result they need to be combined. This is done using the conditional probability table shown as Table 2 to compute the combined belief distribution $D(x)$. $D_\Delta(true)$ represents the probability that Δ is a diagnosis, while $D_\Delta(false)$ represents the probability that it is not. Note that this is different from $B_{D,\Delta}(false)$ as $B_{D,\Delta}(false)$ represents the belief that the evidence does *not* show that Δ is a diagnosis, whereas $D_\Delta(false)$ represents the belief that the evidence *does* show that Δ is *not* a diagnosis. $D_\Delta(unknown)$, represents the belief that we don't know what the evidence shows. The values in Table 2 are chosen such that if $B_{D,\Delta}(x)$ and $B_{IC,\Delta}(x)$ agree that Δ is a diagnosis and not a inverse conflict then $D_\Delta(true) = 1$. However, if they do not agree, then we are confused about what the evidence has shown and $D_\Delta(unknown) = 1$. If neither $B_{D,\Delta}(x)$ nor $B_{IC,\Delta}(x)$ have any information then $D_\Delta(unknown) = 1$.

Table 2. Conditional Probability Table used to combine $B_D(x)$ and $B_{IC}(x)$ into $D(x)$

$$P(D_\Delta(x) \mid B_{D,\Delta}(f_1) = 1 \land B_{IC,\Delta}(f_2) = 1)$$

f_1	f_2	True	False	Unknown
True	True	0.0	0.0	1.0
True	False	1.0	0.0	0.0
True	Unknown	1.0	0.0	0.0
False	True	0.0	1.0	0.0
False	False	0.0	0.0	1.0
False	Unknown	0.0	0.0	1.0
Unknown	True	0.0	1.0	0.0
Unknown	False	0.0	0.0	1.0
Unknown	Unknown	0.0	0.0	1.0

2.4.2 Finding the minimal diagnoses

Definition 7 below, defines a distribution $D_{M_\Delta}(x)$ for each $\Delta \in P(COMPS)$ which represents the belief that the set Δ has the properties of a minimal diagnosis.

Definition 7 *Let $\Delta \in P(COMPS)$.*
Let $\Delta_{C_i} \subset \Delta, i = 1, \ldots, m, \forall i \neq j \ \Delta_{C_i} \neq \Delta_{C_j}$

Let $\Delta_{P_i} \supset \Delta, i = 1, \ldots, n, \forall i \neq j \ \Delta_{P_i} \neq \Delta_{P_j}$.
Define the distribution $\neg D(x)$ such that:

$$\neg D(true) = D(false)$$
$$\neg D(false) = D(true)$$
$$\neg D(unknown) = D(unknown)$$

Define the operator $\odot$ such that $A \odot B$ equals the result of combining A and B using the conditional probability table 3, then:

$$
\begin{aligned}
D_{M_\Delta}(x) = {} & D_\Delta(x) \\
& \odot D_{\Delta_{P_1}}(x) \odot \ldots \odot D_{\Delta_{P_n}}(x) \\
& \odot \neg D_{\Delta_{C_1}}(x) \odot \ldots \odot \neg D_{\Delta_{C_m}}(x)
\end{aligned}
$$

Table 3. Conditional Probability Table used to compute $C = A \odot B$

$$P(C(x) \mid A(f_1) = 1 \land B(f_2) = 1)$$

f_1	f_2	True	False	Unknown
True	True	1.0	0.0	0.0
True	False	0.0	1.0	0.0
True	Unknown	1.0	0.0	0.0
False	True	0.0	1.0	0.0
False	False	0.0	1.0	0.0
False	Unknown	0.0	1.0	0.0
Unknown	True	0.0	0.0	1.0
Unknown	False	0.0	0.0	1.0
Unknown	Unknown	0.0	0.0	1.0

The result is that $D_{M_\Delta}(x)$ is true for sets which have all properties that a minimal diagnosis should have and false or unknown for all other sets. Because $D_M(x)$ is a continuous distribution over the states $\{true, false, unknown\}$, a function is needed which allows the possible diagnoses to be returned to the diagnostician in order from most likely to least, along with a measure of the algorithm's certainty in the result. The following sorting function is suggested as a good balance between certainty in the result and the belief that the set is a minimal diagnosis:

$$D_{M_\Delta}(true) \cdot \left(1 - D_{M_\Delta}(unknown)\right) \qquad (8)$$

Minimal diagnoses can now be returned to the diagnostician in order from the one with the largest value for Equation 8 to the smallest. The probability that a set is a minimal diagnosis is equal to $D_{M_\Delta}(true)/(1 - D_{M_\Delta}(unknown))$ and the certainty in the result defined by $1 - D_{M_\Delta}(unknown)$.

2.5 Complexity considerations

Calling $\text{TP}(\cdot)$ on every subset of $COMPS$ is an exponential undertaking. If the PGDE is run so that the maximum certainty is achieved in the result, every subset of $COMPS$ would need to be tested and the algorithm would indeed be exponential in time. However, a trade-off can be made between certainty and execution time by using some of the criteria listed in Section 2.3.

Maintaining the distributions $B_D(x)$ and $B_{IC}(x)$ is exponential in space if the entire set $P(COMPS)$ is considered. However, for example, we assume that the likelihood of 40 components failing simultaneously in a system of 50 components is negligible. Therefore, the algorithm does not require that the distributions $B_D(x)$ and $B_{IC}(x)$

cover all of $P(COMPS)$, but only up to the level where a reasonable number of simultaneous faults are considered.

As seen in Figure 1 there are four steps to the algorithm which are performed in an iterative fashion: choose next set, call $\mathrm{TP}(\cdot)$, interpret the results and update the beliefs $B_\mathrm{D}(x)$ and $B_\mathrm{IC}(x)$. This algorithm is primarily intended for the diagnosis of complex dynamic systems for which $\mathrm{TP}(\cdot)$ will require a period of simulation in order to test for consistency and so it is assumed that this call will take a significant period of time. Computing the next set to test can be a function of $P(COMPS)$, but it is assumed that the $\mathrm{TP}(\cdot)$ will take the majority of the time. Both the interpretation of the results and the updating of the belief states involve only the supersets and/or subsets of the set under test, which is a relatively small number when compared to the size of $P(COMPS)$. The final two steps of the algorithm do involve the entire set $P(COMPS)$, but as they are not part of the iterative procedure, their effect on the speed of the algorithm is not significant.

3 DIAGNOSIS OF A HYDRAULIC CIRCUIT

Figure 4 shows a schematic for a single degree of freedom hydraulic manipulator used to test the algorithm presented in this paper. The model is made of eight components as seen in Figure 3: the head-side port of the main valve, the rod-side port of the main valve, the cylinder, the manipulator, the rod-side anti-cavitation valve, the head-side anti-cavitation valve, the exit filter and the check valve. The behavior of the components is described by sets of hybrid dynamic equations which can be found in [6] and [5].

The function $\mathrm{TP}((SD_\Delta)^c, OBS)$ was implemented using a modified version of Hybrid Concurrent Constraint programming, or hcc [3]. The set of hybrid dynamic equations $(SD_\Delta)^c$ is passed to the modified hcc, along with OBS which are the time sequences of the sensor values. The system made of $(SD_\Delta)^c$ and OBS will likely be over-constrained and the resulting simulation will contain several discrepancies between measured and simulated values. These residuals (simulated outputs less measured) will also be time sequences which can be compared to a set of residuals recorded during normal operation to generate a consistency measure, μ_Δ. During the experiments, the system was setup in a position control loop with a sinusoidal input signal at a frequency of 0.25Hz. A period of six seconds is recorded, encompassing a single extension and retraction of the manipulator arm. Six experiments were run, each with the arm under a different failure condition which is common in a system such as this [6, 9]. The failures were caused by manual adjustment of the three valves and one friction plate shown in Figure 4. The faults are assumed to be permanent and to have occurred before the measurements are taken. At each iteration the set to be passed to $\mathrm{TP}(\cdot)$ is selected to maximize the expected decrease in $U = \sum_{\Delta \in P(COMPS)} B_{\mathrm{IC},\Delta}(unknown) + B_{\mathrm{D},\Delta}(unknown)$ and the algorithm is stopped when the change in U is less than 1% for more than 10 iterations.

The six failures and the results of fault isolation using the PGDE are as follows. On average, 99.90% of the time taken is spent in simulation during the calls to $\mathrm{TP}(\cdot)$, while only 0.10% is required for the PGDE calculations. For details refer to [5].

- Leak in the hose connecting the valve to the head-side of the cylinder.

  This failure was correctly isolated in all 10 sample runs taking an average of 54.5 seconds.
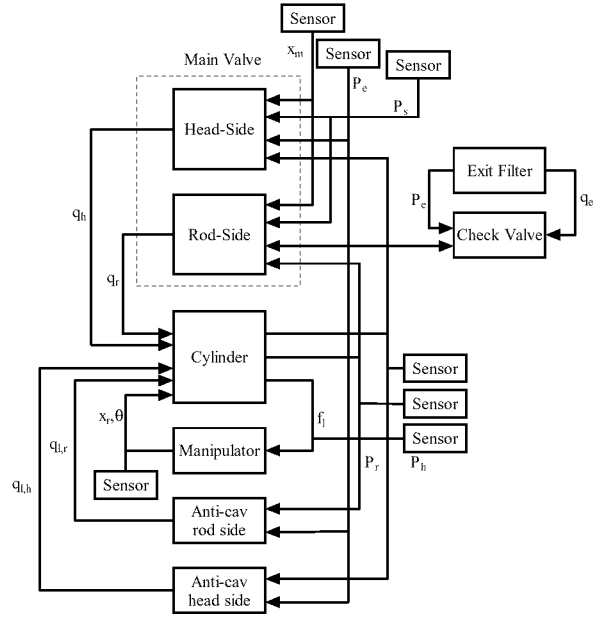- Leak in the hose connecting the valve to the rod-side of the cylinder.

Figure 3. Component model of the hydraulic test bench

This failure was correctly isolated in all 10 sample runs taking an average of 53.1 seconds.
- Partially clogged return filter.

  For two of the five tests run, the filter was returned as the most likely diagnosis, with the rod-side port of the main valve and the rod-side anti-cavitation valves together forming a close second. In the remaining three tests the filter was not returned as a diagnosis by itself, but five diagnoses containing the filter and another component were returned as all being very likely. The average calculation time was 167 seconds.
- Increased friction in manipulator bearing.

  For two of the five tests run, the manipulator was returned as the only likely diagnosis with very high certainty (96%, 100%). In two more of the tests it was returned as one-half of a double fault and in the fifth test the algorithm did not get the correct solution. These calculations took on average 82 seconds to complete.
- Leaks in both hoses connecting the valve to the cylinder.

  In all five tests the four double faults: {rod-side anti-cavitation valve, head-side anti-cavitation valve}, {rod-side anti-cavitation valve, head-side port}, {head-side anti-cavitation valve, rod-side port} and {head-side port, rod-side port} were returned as being equally likely with a high degree of certainty ($\sim$ 85%). For this situation, these are the correct diagnoses as one component on the rod-side and one on the head-side that can account for the leaks is needed to explain this failure. The average calculation time was 140 seconds.
- Partially clogged return filter and a leak in the head-side hose.

  In all five tests the algorithm returned the head-side anti-cavitation valve or port as the only explanation. The filter causes a much smaller effect on the system and so it is difficult to recognize it as faulty when other components are misbehaving. The average calculation time was 61 seconds.
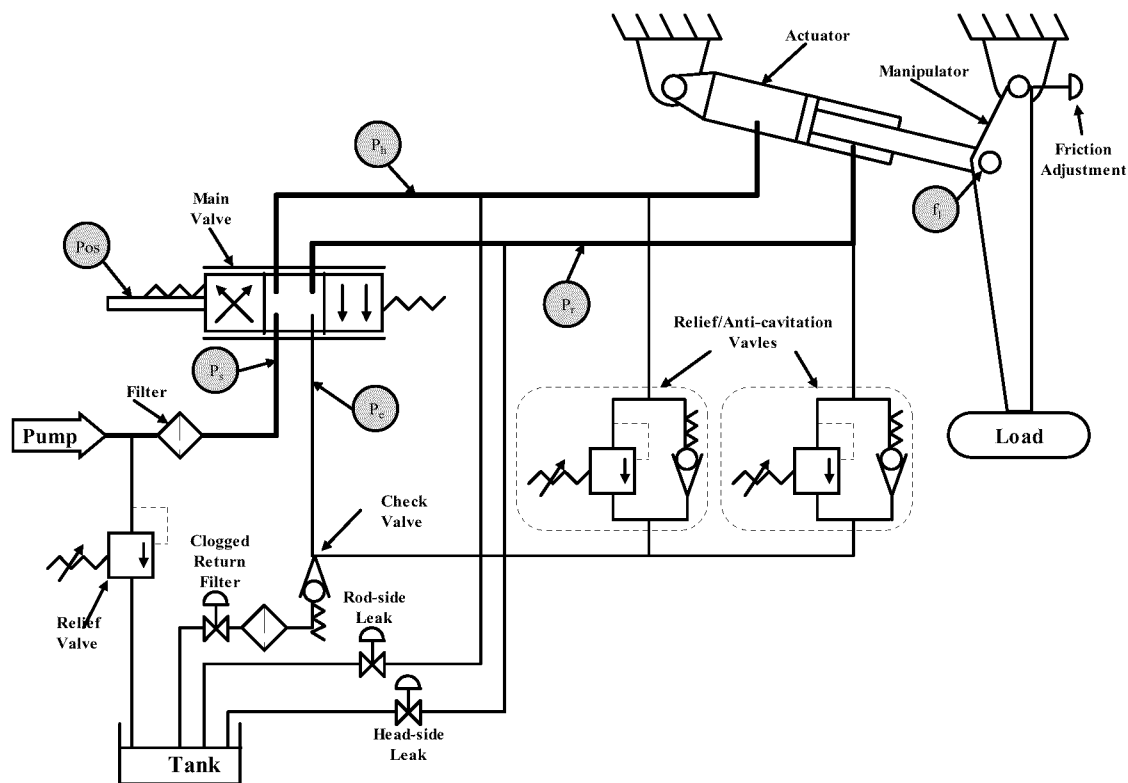
Figure 4. Schematic of experimental test bench

4 CONCLUSIONS

This paper has presented a novel approach to consistency-based diagnosis which allows for the use of any modelling language. The use of continuous distributions representing the belief that each set of components is a diagnosis allows the determination of consistency or inconsistency to be delayed until supporting evidence has been collected and for noise in the simulator, TP($\cdot$), to be handled. The demonstration of this algorithm on a non-trivial physical test bench shows that it can be applied effectively to isolate realistic faults in real artifacts.

REFERENCES

[1] B.Pulido, C.Alonso, C.Llamas, and F.Acebes, 'Consistency-based diagnosis using possible conflicts', in *Proc. of the twelfth Workshop on Principles of Diagnosis*, (2001).

[2] B.Pulido, F.Acebes, and C.Alonso, 'Exploiting knowledge about structure and behaviour in consistency-based diagnosis with fault modes in dynamic systems', in *Proc. of the Ninth Intl. Workshop on Principles of Diagnosis*, (1998).

[3] V. Gupta, R. Jagadeesan, V. Saraswat, and D. Bobrow, 'Computing with continuous change', *Science of Computer Programming*, (1997).

[4] W.C. Hamscher and R. Davis, 'Diagnosing circuits with state: An inherently underconstrained problem', in *Proceedings of the 4th National Conference on Artificial Intelligence*, (1984).

[5] Colin N. Jones, *Consistency-Based Fault Isolation for Hybrid Dynamic Models*, Master's thesis, The University of British Columbia, August 2001.

[6] Masoud Khoshzaban-Zavarehi, *On-Line Condition Monitoring and Fault Diagnosis in Hydraulic System Components using Parameter Estimation and Pattern Classification*, Ph.D. dissertation, University of British Columbia, 1997.

[7] Raymond Reiter, 'A theory of diagnosis from first principles', *Artificial Intelligence*, **32**(1), 57–96, (1987).

[8] S.Narasimhan, F.Zhao, G.Biswas, and E.Hung, 'Fault isolation in hybrid systems combining model based diagnosis and signal processing', in *Proc. of the Intl. Workshop on Hybrid Systems*, (2000).

[9] Xiaodan Sun, *Real-Time Performance Monitoring and Fault Diagnosis of Hydraulic Manipulators*, Master's thesis, The University of British Columbia, March 1995.

[10] Andrew Watkins, *Consistency-Based Diagnosis Using Dynamic Models*, Master's thesis, The University of British Columbia, 1999.